International Journal of Artificial Intelligence in Education
https://doi.org/10.1007/s40593-023-00367-w

OPINION

How to Harness Generative AI to Accelerate Human Learning

W. Lewis Johnson[1]

Abstract

The advent of generative AI has caused both excitement and anxiety among educators. Some school systems have gone so far as to ban it altogether. Generative AI has the potential to transform human learning; but like any new technology, it has both strengths and weaknesses, and adopting it involves some risks. There are risks that generative AI will mislead learners with wrong information, or that learners will use it to do their homework and take tests for them. This article presents some ways to take best advantage of generative AI, while managing and mitigating the risks. It also suggests some uses of generative AI to avoid. These insights are informed by learning science and extensive experience developing AI-enabled learning products. If applied properly, generative AI can dramatically accelerate human learning and do so at scale.

Keywords Generative AI · ChatGPT · Pedagogical agents · Socratic tutors · Simulations

Generative artificial intelligence has the potential to fundamentally transform education and training. Thomas Friedman has argued that we are entering a new Promethean moment (Friedman, 2023), when new tools, ways of thinking, or energy sources are introduced that are such an advance that they change how we work, how we learn–how we do everything. The arrival of generative AI tools such as ChatGPT has caused both anxiety and excitement among educators. On one hand, some have argued that they spell the end of writing assignments, and perhaps even the end of writing as a teachable skill (Herman, 2022). Some school systems, such as New York City schools, have gone so far as to ban ChatGPT from the classroom (Ornstein,

✉ W. Lewis Johnson
ljohnson@alelo.com

[1] Alelo Inc., Los Angeles, USA

Published online: 08 August 2023

2023). Others have suggested that ChatGPT can take on a teaching role itself, for example teaching people how to code in Python(Ahmed, 2022).

I have been developing AIED systems, many incorporating conversational agents, for over thirty years (Johnson & Lester, 2016). I see great potential for generative AI to transform human learning, but only if it is used in the right way. This article presents some ways that generative AI can be used to transform learning. These insights are informed by learning science and extensive experience developing and delivering AI-enabled learning products (Johnson, 2010, 2019, 2021). If applied properly, generative AI can accelerate learning and do so at scale, in ways that were never before possible. I also suggest some uses of generative AI to avoid and how to guard against the risks of generative AI, such as the risk that it will mislead learners with wrong information, or that learners will use it to cheat and do their homework for them.

Strengths and Weaknesses of Generative AI for Education

One of the strengths of generative-AI chatbots is their ability to generate reasonable sounding answers to just about any question, and do so rapidly. If a learner is stuck and is looking for an answer to a particular question, this can be invaluable. If the learner wants further information, the chatbot can elaborate its explanation for them. Generative AI's ability to generate a large number of responses quickly is quite remarkable.

At the same time, generative AI systems have shortcomings that can be significant in an educational context. They sometimes generate answers that are excessively verbose, inconsistent, or just plain wrong. They sometimes "hallucinate" answers by combining various materials found online. Their responses may be biased, reflecting the biases in the data used to train them. The OpenAI website is very clear about these limitations (OpenAI, 2023). This may not be a serious problem for experts and professionals who know the subject area and can recognize inappropriate answers, but novices can be easily misled by plausible-sounding but wrong answers.

Another problem with generative AI is that students can use it to cheat. School administrators are reporting instances of students attempting to pass off writing by ChatGPT as their own work (Klein, 2023). Some educators are concerned that ChatGPT may undercut critical thinking and problem-solving skills (Elliot, 2023). If students simply rely on generative AI to solve problems for them, they miss the learning opportunities that come from solving problems and the critical thinking skills that come from evaluating solutions.

Conversational Avatars Can Avoid the Weaknesses of Generative AI

It is important to keep in mind, however, that the weaknesses and risks of generative AI do not apply to conversational AI in general. Conversational avatars (also known as conversational agents) are a technology that has proven effective in educational settings and in employee upskilling and reskilling. Conversational avatars let learners practice their skills in spoken conversations. Avatars can play a variety of roles in

simulation-based training, acting as simulated customers, co-workers, patients, and coaches. Avatars can also engage learners in Socratic dialogues. Instead of presenting multiple-choice questions and letting learners guess the answers, avatars can ask probing questions so that learners must respond in their own words and apply what they learned. (Khanlabs, 2023). This encourages problem solving and critical thinking (Khan Academy, 2023).

Avatar-based learning tools also benefit learners by encouraging them to practice, in a safe environment where there is no risk of embarrassment from making mistakes. Research shows that for a broad range of learning tasks, learning gains are a function of the number of times learners practice (Koedinger et al., 2023). Each time an avatar asks a question or elicits a response from a learner is a practice opportunity. Therefore, chatbots that answer questions are much less useful as learning tools than avatars that *ask* questions.

Spoken-language avatars are particularly useful because they reveal the learner's degree of mastery of the material. In the context of second language learning, rapid responses are an indicator of cognitive fluency, i.e., a level of mastery at which learners can respond to prompts with a low level of cognitive effort (Segalowitz, 2010). In other contexts rapid responses can indicate low achievement because learners are making quick guessing attempts instead of providing thoughtful answers (Sideridis & Alahmadi, 2022).

Spoken-language avatars are also resistant to cheating. If a learner tries to use ChatGPT to answer an avatar's questions, it will take the learner a long time to respond and it will be very clear that the learner's answers are not their own.

Alelo's avatar-based learning systems have achieved significant results using this approach. In the XPRIZE Rapid Reskilling Competition Alelo's avatar-based training courses upskilled and reskilled community health workers at least twice as fast as conventional training and achieved retention rates that are double that of comparable online courses (Johnson, 2021).

Using Generative AI The Right Way

Generative AI is a powerful technology for AIED systems. But when using generative AI, it is important to avoid and mitigate its weaknesses. Good *prompt engineering* is critical. Prompt engineering is the construction of descriptions of tasks for generative AI to perform. Well-designed prompts greatly reduce the risk of hallucinations and other inappropriate responses. Generative AI systems such as ChatGPT can even be prompted to assess the risk that it will respond inappropriately to a given learner input.

Prompt engineering is therefore a critical core competency for developers of AIED systems. For those who want an introduction I highly recommend the short course by Fulford and Ng (2023). I go even further and argue that prompt engineering is an essential skill for educators and even students, just as coding has been considered an essential skill up to now. It is inevitable that students will seek to take advantage of generative AI, so we should teach them how to use it effectively. In the process they

will develop their critical thinking skills, to evaluate the responses to prompts and consider how to design prompts that yield better responses.

I see the greatest potential for generative AI is not as a learning tool per se but as a *generator of training data for other learning tools*. To develop conversational AI, one needs training data–examples of dialogue utterances and responses. Generative AI can create such examples very quickly. A human expert is still needed to review the generated examples and weed out the "hallucinations" and other inappropriate responses, or redesign the prompts to generate better responses. But reviewing and selecting generated examples, and revising prompts, are much quicker and easier than writing examples from scratch. As learners interact with the avatars the learners' responses are another source of training data. I also recommend giving subject matter experts the option of marking avatar responses that they consider inappropriate or incorrect, since that can help identify poorly engineered prompts and inform the retraining process. Conversational avatars developed using this approach are more likely to interact appropriately with learners than chatbots trained on unfiltered Internet data.

Alelo is putting these techniques to practice in its own conversational AI systems. Soon after the emergence of ChatGPT we started incorporating generative AI into our development pipeline. We use large language models with carefully engineered prompts as a component technology, alongside our own custom natural language understanding models that are trained on learner data. This gives us the flexibility to use generative AI for most common learning tasks and other technologies for more specialized learning tasks (e.g., understanding the errorful language of second language learners, or training workers to employ an organization's specific communication strategies when engaging with customers).

Looking ahead, I see this approach as a way for instructional designers and subject matter experts to create their own instructional avatars. Instead of laboriously scripting avatar responses, they can rely on generative AI to generate candidate responses, and select the responses that they prefer.

Assuming a Support Role Mitigates the Risks of Generative AI

Generative AI technology continues to evolve very rapidly. GPT-4 and other tools have already emerged as successors to GPT-3. Yet meanwhile many tech leaders have called for a moratorium on the development of the most advanced AI systems (Future of Life Institute, 2023) so that their potential risks can be mitigated. AI-based tools that are informed by learning science and designed to promote learning do not pose the same risks. I believe that they will continue to have an advantage over general-purpose question-answering chatbots, even as the technology continues to develop.

As a final note: After I wrote this article I asked ChatGPT to write its own article on this topic. It produced a reasonable and coherent explanation of the potential benefits and risks of generative AI, and it made a suggestion (the risk of bias) that I thought was good and I included it in this article. Of course I was aware of the risk of bias, but neglected to mention it when I wrote the first draft. But overall I found ChatGPT's writing to be bland, not particularly insightful, and lacking in specific

recommendations for AIED developers. Perhaps through careful prompt engineering I might have been able to get ChatGPT to generate a better result. But overall, I am glad that I chose to write this article myself, rather than rely on ChatGPT to write it for me.

References

Ahmed, S. (2022). Learn Python with ChatGPT. The Developer Space, December 26, 2022.

Elliot, L. (2023). Those schools banning access to generative AI ChatGPT are not moving the needle and are missing the boat, says AI ethics and AI law. Forbes, Jan. 20, 2023.

Friedman, T. L. (2023). Our new Promethian moment. The New York Times, March 21, 2023.

Fulford, I., & Ng, A. (2023). ChatGPT prompt engineering for developers. Retrieved May 31, 2023 from https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/.

Future of Life Institute (2023). Pause giant AI experiments; An open letter. Retrieved May 31, 2023 from https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

Herman, D. (2022, December 9). The end of high-school English. *The Atlantic*.

Johnson, W. L. (2010). Serious use of a serious game for language learning. *Int J of Artificial Intelligence in Ed* 20.

Johnson, W. L. (2019). Data-Driven Development and evaluation of Enskill English. *Intl Journal of AI in Education*, *29*, 425–457.

Johnson, W. L. (2021). Case study: Rapid training of community health workers using Enskill®. *International Journal of Advanced Corporate Learning*, *14*(2), 44–49.

Johnson, W. L., & Lester, J. C. (2016). Face-to-face Interaction with Pedagogical Agents, twenty years later. *Int J AI in Education*, *26*(1), 25–36.

Khan Academy (2023). Introducing… Khanamigo! Retrieved May 31, 2023 from https://support.khanacademy.org/hc/en-us/community/posts/13992414612877-Introducing-Khanmigo.

Klein, A. (2023). ChatGPT Cheating: What to do when it happens. EducationWeek, Feb. 21, 2023.

Koedinger, K., Carvalhu, P. F., Liu, R., & McLaughlin, E. A. (2023). An astonishing regularity in student learning rate. *PNAS* 120:13, Mar 30, 2023. https://doi.org/10.1073/pnas.2221311120.

OpenAI (2023). Introducing ChatGPT. Retrieved May 31, 2023 from https://openai.com/blog/chatgpt.

Ornstein, V. (2023). ChatGPT: What it is and why it is banned in NYC schools. New York Family, March 14, 2023.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.

Sideridis, G., & Alahmadi, M. T. S. (2022). The role of response times on the measurement of mental ability. *Front. Psychol* 13, June 17, 2022. https://doi.org/10.3389/fpsyg.2022.892317.